

Common Time Variation of Parameters in Reduced-Form Macroeconomic Models

Dalibor Stevanovic *

Previous version: May 2014

This version: March 2015

Abstract

Standard time varying parameter (TVP) models usually assume independent stochastic processes for all TVPs. In this paper, I show that the number of underlying sources of parameters' time variation is likely to be small, and provide empirical evidence for factor structure among TVPs of popular macroeconomic models. To test for the presence of, and estimate low dimension sources of time variation in parameters, I develop the factor time varying parameter (Factor-TVP) framework and apply it to Primiceri (2005) monetary TVP-VAR model. I find that one factor explains most of the variability in VAR coefficients, while the stochastic volatility parameters vary independently. Including post-“Great Recession” data causes an important change within VAR coefficients and the procedure suggests two factors. The roots of variability in the VAR parameters are likely to come from the financial markets and the real sector. The TVP factors have predictive power for a large number of output, investment and employment series as well as for the term structure of interest rates.

JEL Classification: C32, E17, E32.

Keywords: time varying parameter model, factor analysis, VAR.

*Département des sciences économiques, Université du Québec à Montréal. 315, Ste-Catherine Est, Montréal, QC, H2X 3X2. (dstevanovic.econ@gmail.com)

I would like to thank Jean Boivin, Frank Diebold, Jean-Marie Dufour, Jesus Fernandez-Villaverde, and Frank Schorfheide for valuable discussions and comments. A previous version of this paper circulated as “Common Sources of Parameter Instability in Macroeconomic Models: A Factor-TVP Approach”. The author acknowledges financial support from the Fonds de recherche sur la société et la culture (Québec).

Contents

1	Introduction	2
2	Examples of reduced-rank parameters instability	4
2.1	Vector autoregressive models	4
2.2	General equilibrium models	5
3	Factor representation of time-varying parameters' process	7
3.1	Linear approximation	7
3.2	Dimension reduction	8
4	Econometric framework	8
4.1	Factor-TVP model	8
4.2	Identification	9
4.3	Estimation	10
4.3.1	Estimation of the number of factors	11
4.4	Simulations	12
5	Empirical evidence on common sources of parameters instability	14
5.1	TVP-VAR model	14
5.1.1	Detecting common parameters' instability	15
5.1.2	Results from the <i>Factor</i> -TVP VAR	17
5.1.3	Further analysis	19
5.2	FAVAR model	21
6	Conclusion	22

1 Introduction

It is likely that the behavior of economic agents and environment vary over time: monetary policy authority changes its strategy, economic shocks become more or less volatile, etc. This implies time instability in potentially all parameters in reduced-form representations of structural models. For instance, Fernandez-Villaverde and Rubio-Ramirez (2008) find that some parameters in dynamic stochastic general equilibrium (DSGE) models are time-varying. Using empirical univariate and bivariate autoregressive models, Stock and Watson (1996) find widespread instability in a large number of U.S. macroeconomic series. In addition, several studies, while investigating the causes of the Great Moderation, have assumed or tested structural changes such as shifts in monetary policy rule parameters and/or time changes in volatility of shocks (see Boivin and Giannoni (2006), and Stock and Watson (2002, 2003), among others). The time-varying parameter VAR models (TVP-VAR hereafter) were also used by Cogley and Sargent (2005), Boivin (2005), Primiceri (2005) and Sims and Zha (2006) to investigate the time instability of policy functions and the stochastic volatility of structural shocks. On the other hand, a number of studies using DSGE models assumed some of the structural parameters to be time-varying and/or impose stochastic volatilities [see e.g. Justiniano and Primiceri (2008), Fernandez-Villaverde et al. (2010) and Ravenna (2010)]. Finally, Inoue and Rossi (2011) used a sequential testing procedure to identify which structural parameters in both VAR and DSGE models are time-varying.

Most of the studies using reduced-form models assumed the same number of sources of time variations as the number of time-varying parameters. Due to the computational difficulty, independent stochastic processes are imposed for all coefficients.¹ However, it is likely that this time variability presents commonalities. The intuition is that only a small number of structural relationships vary over time, inducing instability in all coefficients of reduced-form models, where the correlation structure between TVPs is mainly explained by the common component. An exception is Cogley and Sargent (2005) who find that main time instability within the VAR coefficients can be explained by only a few components. In addition, Canova and Ciccarelli (2009) impose a reduced-rank time variation within the estimation of a multicountry panel VAR while Carriero, Clark and Marcellino (2012) impose common factor stochastic volatility in a large Bayesian VAR.

In this paper, I provide new evidence that the number of common sources of parameters time variation in widely-used empirical macroeconomic models is very small. There are several contributions to the literature. I first show that parameters instability in the structural model is likely to imply time variation in all (or at least a subset of) coefficients in the reduced-form model. This is explicitly demonstrated for structural VARs, and VAR(MA) representations of DSGE models. I

¹Primiceri (2005) relaxes this hypothesis by letting some VAR time-varying coefficients have correlated error terms.

demonstrate how the factor representation of time-varying parameters can be obtained, and present the factor time varying parameter model (Factor-TVP) that takes into account the commonalities within the coefficients. From a practical point of view, the main advantages of the model are: the correlation structure between the TVPs is unrestricted, and only a small number of states must be filtered. The approach is applied to a standard 3-variable VAR model from Primiceri (2005), and to a factor-augmented VAR (FAVAR) model from Boivin, Giannoni, and Stevanovic (2013).

The first objective is to detect the presence of the factor structure, and test its dimensionality in time-varying VAR coefficients, keeping the volatility of shocks constant. Performing the two-step recursive and likelihood-based procedures, I find that a small number of common shocks explain the most of instability in VAR coefficients. Then, the Factor-TVP model with one latent component is estimated by maximum likelihood. The underlying factor is very persistent, which is in line with the strong collinearity in TVPs found after the first stage estimation in the two-step likelihood procedure. Moreover, it is highly correlated with the unemployment rate, and moderately related to inflation and interest rate.

When applied to TVP-VAR model with stochastic volatility from Primiceri (2005), the variability in VAR coefficients is mostly explained by only one factor, while the stochastic volatility part is explained by two additional factors. In particular, when the structural version of the VAR model is obtained by the Choleski decomposition of the residuals time-varying covariance matrix, I find that the contemporaneous relations coefficients are explained by the second factor, while the variances of structural shocks are explained by the third factor.

I repeat the same exercise with the data updated to 2013Q4. The idea of using the Factor-TVP model is to capture common breaks, and the recent financial crisis is a good example of such an episode. Indeed, the tests suggest one more factor, and their interpretation is changed. The estimated TVPs present an important behavioral change after 2007, indicating the presence of a structural break. The correlation structure between factors and observable variables is such that the first factor is highly positively correlated to the inflation rate, and moderately related to the interest and unemployment rates. The second factor presents a similar pattern, but with correlation coefficients around 0.3. Contrary to pre-crisis data, a majority of TVPs are largely explained by the factors. In particular, the inflation equation coefficients seem to be time-varying with the two factors, and this is also the case for about half of parameters in other VAR equations.

To improve on the interpretation of TVP factors and to understand more deeply their roots and implications for empirical research, I face them against a large macroeconomic panel containing 141 time series. The findings are following: (i) the roots of variability in the VAR parameters come from the financial markets and from the real sector of economic activity, (ii) the TVP factors are Granger-caused by the macroeconomic fundamentals, as estimated by static factors from the large

data set, (iii) the TVP factors have significant predictive power, beyond the information contained in data factors, for a great number of output, investment and employment series as well as for term structure of interest rates, (iv) this predictive content increases with the forecasting horizon. This evidence suggests that more attention should be given to the link between the financial and real sectors of economy within structural models.²

To complete the empirical exercise, the time-varying FAVAR model is estimated by the two-step recursive procedure. I find that four dynamic factors explain most of the commonalities between almost 700 TVPs. Again, the stochastic volatility coefficients' instability seems to have different common factors than the regression parameters. In the rest of the paper, I first provide examples of common sources of parameters' instability in macroeconomic models. Section 3 shows how the factor representation of time-varying parameters process is obtained, and the Section 4 presents the Factor-TVP model. The main results are presented in Section 5 and Section 6 concludes.

2 Examples of reduced-rank parameters instability

In the following, I show two examples where the low number of common sources of parameter instability is a plausible hypothesis. The intuition is that only a small number of structural relationships vary over time, which implies that possibly all the coefficients in reduced-form models are unstable.

2.1 Vector autoregressive models

Suppose the two-dimensional stochastic process $Y_t = (y_{1t} \ y_{2t})'$ has a structural VAR(1) representation:

$$\begin{bmatrix} a_{0,11} & 0 \\ a_{0,21} & a_{0,22} \end{bmatrix} \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{1,11} & a_{1,12} \\ a_{1,21} & a_{1,22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

with $V(\varepsilon_t) = I$. Then, the reduced-form representation is:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \frac{a_{1,11}}{a_{0,11}} & \frac{a_{1,12}}{a_{0,11}} \\ \frac{a_{1,21}}{a_{0,21}} - \frac{a_{0,21}a_{1,11}}{a_{0,11}a_{0,22}} & \frac{a_{1,22}}{a_{0,22}} - \frac{a_{0,21}a_{1,12}}{a_{0,11}a_{0,22}} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \frac{1}{a_{0,11}} & 0 \\ -\frac{a_{0,21}}{a_{0,11}a_{0,22}} & \frac{1}{a_{0,22}} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

or in matrix notation

$$Y_t = BY_{t-1} + u_t$$

²In fact, Philippon (2012) provides evidence on the growing importance and the cost of financial intermediation in the U.S. in the last 130 years. Barattieri, Eden and Stevanovic (2014) suggest that the increase in financial sector interconnectedness (a decrease of direct connectedness with the real sector) may have dampened the sensitivity of the real variables to monetary shocks. There is also a growing literature on shadow banking and its impact on real economy, see Poszar et al. (2012).

with

$$V(u_t) = \Omega = \begin{bmatrix} \frac{1}{(a_{0,11})^2} & - \\ -\frac{a_{0,21}}{a_{0,11}a_{0,22}} & \frac{1}{a_{0,22}} + \frac{(a_{0,21})^2}{(a_{0,11})^2(a_{0,22})^2} \end{bmatrix}$$

Let $a_{0,11}$ be time varying, i.e. $a_{0,11}^t \sim G$ where G is some distribution. Then, all reduced-form coefficients in B and Ω are time varying such that:

- Distributions of reduced form parameters depend in generally nonlinear way on distribution of time-varying structural parameters.
- Variability of elements in B_t and Ω_t is of reduced-rank.
- All the correlation structure between the elements of B_t and Ω_t is explained by the time-varying parameters in structural form.

\Rightarrow All these point to a (non)linear factor structure.

2.2 General equilibrium models

It is well known that the linearized solution of a DSGE model has a vector-autoregressive moving-average (VARMA) form in endogenous variables (see Fernandez-Villaverde et al. (2007) and Ravenna (2007)), but the coefficients of the corresponding VARMA model are nonlinear combinations of parameters in the structural model. As such, the parameters instability in the structural model is likely to imply time variation in all coefficients of the VARMA representations. The evidence on time-varying parameters in macroeconomic general equilibrium models is growing, see Fernandez-Villaverde and Rubio-Ramirez (2008) and Inoue and Rossi (2011). In the following, I show how this may happen in a simple general equilibrium model.

Consider a forward-looking model with backward component

$$\pi_t = \frac{\beta}{1 + \alpha\beta} E_t(\pi_{t+1}) + \frac{\alpha}{1 + \alpha\beta} \pi_{t-1} + \kappa y_t + \varepsilon_{\pi,t} \quad (1)$$

$$y_t = \gamma E_t(y_{t+1}) + (1 - \gamma)y_{t-1} - \frac{1}{\sigma}(R_t - E_t(\pi_{t+1})) + \varepsilon_{y,t} \quad (2)$$

$$R_t = \rho R_{t-1} + (1 - \rho)(\phi_\pi \pi_t + \phi_y y_t) + \varepsilon_{R,t} \quad (3)$$

$$\varepsilon_{i,t} = \rho_i \varepsilon_{i,t-1} + \epsilon_{i,t} \quad (4)$$

with $\epsilon_{i,t} \sim N(0, \sigma_i^2)$, $i = \pi, y, R$. This model has a general solution:

$$g_t = G_1 g_{t-1} + C \epsilon_t \quad (5)$$

where $g_t = [R_t, \pi_t, y_t, \xi_t^\pi, \xi_t^y, e_{R,t}, e_{\pi,t}, e_{y,t}]'$, $\xi_t^\pi \equiv E_t(\pi_{t+1})$, $\xi_t^y \equiv E_t(y_{t+1})$, $\epsilon_t = [\epsilon_{R,t}, \epsilon_{\pi,t}, \epsilon_{y,t}]'$, G_1 is 8×8 and C is the 8×3 impact matrix.

Let $X_t = [R_t, \pi_t, y_t]'$, $\Xi_t = [\xi_t^\pi, \xi_t^y]'$ and $\varepsilon_t = [\varepsilon_{R,t}, \varepsilon_{\pi,t}, \varepsilon_{y,t}]'$. Then we can rewrite (5)³

$$\begin{bmatrix} A_{11}(L) & A_{12}(L) & A_{13}(L) \\ A_{21}(L) & A_{22}(L) & A_{11}(L) \\ A_{31}(L) & A_{32}(L) & A_{33}(L) \end{bmatrix} \begin{bmatrix} X_t \\ \Xi_t \\ \varepsilon_t \end{bmatrix} = \begin{bmatrix} C_{11} \\ C_{21} \\ C_{31} \end{bmatrix} \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{\pi,t} \\ \epsilon_{y,t} \end{bmatrix}$$

From the third system of equations in previous representation, assume $A_{31}(L)$ and $A_{32}(L)$ are zero and solve for ε_t to get

$$\varepsilon_t = A_{33}(L)^{-1}C_{31}\epsilon_t. \quad (6)$$

Solve for Ξ_t in the second system and replace for e_t using (6)

$$\Xi_t = -A_{22}(L)^{-1}A_{21}(L)X_t - A_{22}(L)^{-1}A_{23}(L)A_{33}(L)^{-1}C_{31}\epsilon_t + A_{22}(L)^{-1}C_{21}\varepsilon_t. \quad (7)$$

Finally, solve for X_t and use (6-7) to get

$$\begin{aligned} [A_{11}(L) - A_{12}(L)A_{22}(L)^{-1}A_{21}(L)]X_t &= [C_{11} + A_{12}(L)A_{22}(L)^{-1}A_{23}(L)A_{33}(L)^{-1}C_{31} - A_{13}(L)A_{22}(L)^{-1}C_{31} \\ &\quad - A_{12}(L)A_{22}(L)^{-1}C_{21}]\epsilon_t. \end{aligned} \quad (8)$$

Note that this VARMA(∞, ∞) process can be written in a finite order VARMA(p, q) representation by observing that $A_{ij}(L)^{-1} = \frac{\text{adjoint}(A_{ij}(L))}{\det(A_{ij}(L))}$. Since the coefficients in $A(L)$ are nonlinear combinations of the structural parameters, if q of them are unstable then all elements of $A(L)$ are time varying but with only q sources of variability.

These examples have an important implication for the counterfactual analysis that is widely used in the empirical macroeconomic literature. It usually consists of combining the coefficients of reduced-form model from one regime with those from the second regime. However, a single regime change in the structural model may imply time variation and/or stochastic volatility in all the coefficients of the reduced-form representation, even in the structural VAR. Moreover, the mapping to the structural instability is probably highly nonlinear. Hence, there is no reason that one time change in some structural parameters in a particular equation corresponds to one time change in the corresponding VAR equation coefficients. Thus, supposing this correspondence to be true can be very misleading and makes the resulting counterfactual inappropriate.

³Generally $A_{31}(L)$ and $A_{32}(L)$ are zero, while the coefficients in $A_{12}(L)$ are all zero only in determinacy case.

3 Factor representation of time-varying parameters' process

3.1 Linear approximation

Let α_t be a q -dimensional vector of time-varying parameters in the structural model, whose dynamic process is characterized by some distribution \mathcal{G} . Let β_t be a K -dimensional vector of (potentially all) time-varying *reduced-form* coefficients such that

$$\beta_t = \mathcal{F}(\alpha_t; \gamma) \quad (9)$$

where γ is a vector of m constant coefficients, and \mathcal{F} is some functional form that links reduced-form to structural parameters.

In general, β_t is a nonlinear function of structural parameters. If the number of TVPs in the structural model is less than the number of coefficients in β_t , i.e. $q < k$, then $\mathbf{rank}(\Sigma) = q$, $\Sigma = \mathbf{V}(\beta_t)$. Suppose that function \mathcal{F} in (9) is twice continuously differentiable, then by Taylor's theorem

$$\beta_t = \mathcal{L}(\beta_t) + error \quad (10)$$

where $\mathcal{L}(\beta_t)$ is the linear approximation of β_t around $\bar{\alpha}$:

$$\mathcal{L}(\beta_t) = \mathcal{F}(\bar{\alpha}; \gamma) + \mathcal{F}'(\bar{\alpha}; \gamma)(\alpha_t - \bar{\alpha}).$$

Defining $\mu = \mathcal{F}(\bar{\alpha}; \gamma)$, $\lambda = \mathcal{F}'(\bar{\alpha}; \gamma)$, $f_t = (\alpha_t - \bar{\alpha})$ and e_t as the approximation error, we have the following factor structure:

$$\beta_t - \mu = \lambda f_t + e_t \quad (11)$$

If β_t is demeaned, then $\mu = 0$ and we obtain the usual factor model:

$$\beta_t = \lambda f_t + e_t$$

with $f_t \sim \bar{\mathcal{G}}$, where $\bar{\mathcal{G}}$ is the distribution of $(\alpha_t - \bar{\alpha})$, and with dispersion matrices $\mathbf{D}(e_t) = \Psi$ and $\mathbf{D}(f_t) = \Phi$ and

$$\mathbf{D}(\beta_t) = \lambda \Phi \lambda' + \Psi$$

with Ψ diagonal and at least q^2 identifying restrictions.

3.2 Dimension reduction

Another way to motivate the factor representation of time-varying parameters is the dimension reduction argument. Since the standard estimation methods are computationally cumbersome when the number of TVPs is large, it may be of practical interest to reduce the dimensionality problem by imposing the factor structure. As in standard applications of principal component analysis, the idea is to approximate a large number of TVPs by few factors. Suppose we wish to replace the K -dimensional random variable β_t with its q , $q < K$, linear functions without much loss of information. How are the best q linear function to be chosen? It can be restated as a nonlinear least square problem

$$\beta_{it} = (B'_1\beta_t)\delta_{i1} + \dots + (B'_K\beta_t)\delta_{iK} + e_{it}, \quad i = 1, \dots, K, \quad t = 1, \dots, T. \quad (12)$$

Rao (1973) shows that $B'\beta_t$ should be constructed as the first q principal components of β_t . Hence, rewriting (12) in matrix form and defining $A = \delta$, and $F = B'\beta$, we get the linear factor representation:

$$\beta = AF + e.$$

4 Econometric framework

4.1 Factor-TVP model

The simplest example to illustrate the method is the linear TVP regression:

$$y_t = x'_t\beta_t + w_t \quad (13)$$

where y_t is a scalar, x_t is a $K \times 1$ vector of explanatory variables, β_t contains K time-varying parameters and w_t is a homoscedastic white noise with $\mathbf{V}(w_t) = R$. There are several ways to specify the time variation of β_t : discrete or stochastic breaks, stochastic continuous processes (random walk and ARIMA). In standard applications of TVP models in macroeconomics, β_t is usually modeled either as a Markov switching process, which includes discrete break as special case, or as a collection of univariate AR(1) processes, which embodies the random walk if the autoregressive coefficient is fixed to unity. Here, we concentrate only on continuous stochastic processes of β_t .⁴ Following intuition from the previous section, the link between β_t and its underlying factors may be approximated as a linear factor model:

⁴Nyblom(1989) pointed out that both discrete break and random walk models are martingale processes and are special cases of $\beta_t = \beta_{t-1} + \omega_t$, with $\mathbf{E}(\omega_t) = 0$. Moreover, the continuous TVP model is then an approximation of the discrete break model.

$$\beta_t = \lambda f_t + e_t \tag{14}$$

$$f_t = \rho f_{t-1} + v_t \tag{15}$$

where f_t is a $(q \times 1)$ vector of latent factors, λ is a $(K \times q)$ matrix of factor loadings, ρ is such that f_t is stationary of martingale, e_t and v_t are white noise processes with $E(e_t e_t') = M$ and $E(v_t v_t') = Q$. Finally, it is assumed that w_t , e_t , and v_t are uncorrelated.

To summarize, the Factor-TVP model is given by the following expressions:

$$y_t = x_t' \beta_t + w_t \tag{16}$$

$$\beta_t = \lambda \beta_{t-1} + e_t \tag{17}$$

$$f_{t+1} = \rho f_t + v_{t+1} \tag{18}$$

4.2 Identification

As written above, the Factor-TVP model is not identified. Suppose that $\tilde{\lambda}$ and \tilde{f}_t are a solution to the estimation problem. However, this solution is not unique since we could define $\hat{\lambda} = \tilde{\lambda}H$ and $\hat{f}_t = H^{-1}\tilde{f}_t$, where H is a $q \times q$ nonsingular matrix, which could also satisfy the model's equations. Then, observing y_t and x_t is not enough to distinguish between these two solutions, and a set of normalization restrictions is necessary.

There are several ways to achieve identification in the factor model. Recall that we have to impose q^2 restrictions. In principal component analysis, this is done by rotating factors such that they are orthogonal and with unit variance. This gives $q(q + 1)/2$ restrictions. The remaining restrictions are obtained by assuming that $\lambda'\lambda$ is diagonal.

However, this does not remove the indeterminacy associated with rotation, orthogonal transformation, or sign changes of f_t . The main criticism to this normalization is that the factors are not interpretable. If the objective is to estimate the space spanned by factors and then use this information to improve forecasting of observable series, this normalization is adequate. However, as pointed out in Yalcin and Amemiya (2001), if one is interested in a causal use of factor analysis, this parametrization can be problematic.

Another parametrization of (17) is the errors-in-variable representation. The identification is achieved by imposing a $(q \times q)$ identity matrix on q rows of λ :

$$\tilde{\lambda} = \begin{bmatrix} I \\ \lambda \end{bmatrix}.$$

This is equivalent to re-write (18) as

$$\begin{aligned}\beta_{1,t} &= f_t + e_{1,t} \\ \beta_{2,t} &= \lambda f_t + e_{2,t}\end{aligned}$$

which implies that f_t is measured with error by q elements of $\beta_{1,t}$.

The main advantage of the errors-in-variable representation is that only q^2 loadings are restricted while the factors distribution is left unrestricted. This can be of interest if the research goal is to study the causal relation between factors and time-varying parameters. For example, if one seeks to learn where the instability in VAR parameters is coming from, it is important to not restrict the factors distribution. However, the major problem with this parametrization is that one must arbitrarily choose which q time-varying parameters belong to $\beta_{1,t}$.

4.3 Estimation

The strength of the Factor-TVP model is that instead of filtering and tracking $K + N(N + 1)/2$ states, we only need to filter q states, and q is generally much less than $K + N(N + 1)/2$. Moreover, much less covariance coefficients must be estimated. Consider an N -dimensional TVP-VAR(1) model with stochastic volatilities. If no factor structure is imposed one must filter $N + N(N + 1)/2$ states, but also estimate the same number of TVP variances. This is the most enthusiastic case in which all coefficients are modeled as random walks with a diagonal covariance matrix. If one is interested in more flexible specifications the number of parameters to estimate will explode.

On the other hand, imposing a factor structure allows for any type of correlation between TVPs, while keeping the dimensionality of the estimation problem relatively low. The number of TVP covariance elements to be estimated is at most $q(q + 1)/2$, but the number of factor loadings is at most $(N + N(N + 1)/2)q$ in case of orthogonal factors and $(N + N(N + 1)/2)q - q^2$ if errors-in-variable representation.

In this paper, I consider three estimation procedures. The first consists of complete characterization of the system (16)-(18) in a parametric world and simultaneous estimation of the system using a likelihood-based method. In addition to elements of β_t , it is possible to estimate the stochastic volatility so that the covariance matrix of residuals in (16) is heteroskedastic. If no stochastic volatility, and if the number of parameters is not too large, the model can be estimated by ML where the likelihood is calculated using the Kalman filter. Otherwise, Bayesian methods are needed⁵.

⁵An ongoing research project is studying the estimation of the Factor-TVP-VAR model with stochastic volatility within the Bayesian framework, see Amir-Ahmadi and Stevanovic (2015).

Instead of the simultaneous parametric methodology, one can estimate (16)-(18) by recursive least squares. On one hand, the recursive methods may be inefficient if the true restrictions are not explicitly imposed and the estimates may contain a large sampling error, but they are more robust to misspecification.⁶ Also, they are computationally easy to apply contrary to the simultaneous likelihood methods that are often numerically unstable. The procedure works in two steps:

Step 1: Estimate TVPs by recursive procedure with rolling or expanding window.

Step 2: Perform factor analysis on estimated TVPs.

In the first step, several choices must be made: size of the time window, rolling or expanding window. Then, once we have the set of estimated TVPs, an appropriate factor analysis must be applied: dynamic or static factor model, the number of factors, identification and the estimation method (principal components or MLE).

Finally, the third method combines the two previous procedures. In the first step, the TVPs are estimated within the standard TVP framework without imposing the factor structure, and then the factor model is fitted on estimated TVPs. According to simulation results presented in following section, this method provides accurate estimates of the number of common shocks, and of the factors' dynamic process, but is naturally less efficient than the first method since the true restrictions are not imposed.

4.3.1 Estimation of the number of factors

The most difficult part in the factor analysis is to test for the number of factors. I will use *scree tests*, i.e. plot eigenvalues of the correlation matrix of data; *trace test*, to see the percentage of variance explained by factors; Bai and Ng (2007) test for the number of dynamic factors, Horn (1965) parallel tests and Velicer (1976) MAP test.⁷ As pointed out in Velicer et al. (2000), the scree test is good as an adjunct method since it is not a statistical procedure and the trace test is not recommended since it is not robust to the presence of irrelevant components.⁸ The parallel test, that is in fact a simulation-based Monte Carlo test, is one of the most recommended but it is computationally difficult if there is a large number of TVPs. The MAP test (Velicer (1976) and Velicer et al. (2000)) is also a widely-used test in psychometric applications of factor analysis, but

⁶The results in Stock and Watson (1996) suggest that time-varying models have limited success in exploiting this instability to improve upon fixed-parameter or recursive autoregressive forecasts. In addition, Edlund and Sogaards (1993) find that recursive methods are useful to model business cycles in the context of time-varying parameters.

⁷Given the sample size and sampling error when TVPs are estimated, I do not report results from Bai and Ng (2002) for number of static factors since these performed quite poorly in simulations. This was expected since the cross-section size is small while Bai and Ng (2002) information criteria asymptotics require a large number of elements.

⁸Adding irrelevant series to a data set constructed from factor model will decrease the trace ratio. This is important to note since we do not know a priori how many parameters are time varying due to common shocks.

it tends to overestimate the number of factors (especially in simulations where a lot of sampling error is present). Finally, likelihood ratio tests are known to overestimate the number of factors.

4.4 Simulations

In this section I check the performance of estimation procedures in simulations. Several Factor-TVP models are considered. No stochastic volatility is assumed so only regression coefficients are time varying.

- Model 1: Univariate Factor-TVP Regression

$$y_t = x_t' \beta_t + w_t \quad (19)$$

$$\beta_t = \lambda f_t + \sigma_\beta e_t \quad (20)$$

$$f_{t+1} = \rho f_t + v_{t+1} \quad (21)$$

where y_t is a scalar, x_t is $(K \times 1)$ and f_t contains q elements. The error terms are all drawn independently from a standardized Normal distribution. σ_β is the variance of e_t , and in particular when it is equal to 0, we have a case where the time-varying parameters are exact linear combinations of underlying common shocks. Note that x_t includes the constant.

- Model 2: Multivariate Factor-TVP Regression

$$y_t = c_t + b_t x_t + w_t$$

where y_t is $(N \times 1)$ and x_t contains K independent variables. Let $X_t = I_N \otimes [1, x_t']$ and $\beta_t = [\text{vec}(c_t)', \text{vec}(b_t)']'$ contains all the $K(N + 1)$ time-varying coefficients. Then the Factor-TVP is

$$y_t = X_t' \beta_t + w_t \quad (22)$$

with (20)-(21) describing the TVPs dynamics.

- Model 3: Factor-TVP VAR

I consider only VAR(1) specification since it is always possible to rewrite the VAR(p) in its companion form:

$$y_t = c_t + b_t y_{t-1} + w_t.$$

The Factor-TVP representation is the same as in the previous multivariate case, but the number of elements in β_t is $N(N + 1)$.⁹

⁹The simulation of TVP-VAR model is tricky. To avoid explosive roots, for each t , I decompose b_t using singular

The simulation results are presented in Table 1. I generate quite persistent movements in β_t by fixing factors autoregressive coefficients close to one. The objective is to mimic potential structural breaks in economy that are likely to be discrete or slow moving. The two-step rolling OLS method performs quite well to detect the presence of commonalities within TVPs.¹⁰ It tends to overestimate the number of factors in case of multivariate models. The intuition is that the first principal component is extracted such that the maximal amount of variance in $\hat{\theta}_t$ is explained while the recursive estimation of coefficients introduces some sampling uncertainty and increases artificially the variability in TVPs. As the number of factors augments, Parallel tests tend to pick the true value of q , while Bai-Ng specifications may even underestimate it. The persistence in dynamic process of f_t is usually overestimated while the estimated factor explains a sizeable proportion of variation in the true factor. Hence, this method seems to be a good starting point to check for the presence of a factor structure on TVPs. This is important in practice given the complexity and sensitivity of standard time-varying parameter estimation procedures.

The two-step MLE method, where the first step consists of a standard TVP model (without factor structure), produce accurate estimates of q . In addition, the MLE factors span reasonably well the space generated by the true factors, both in terms of R^2 and root mean squared error (RMSE). However, it tends to overestimate the persistence in factors dynamics and underestimate their variance. Naturally, the simultaneous MLE of Factor-TVP gives the best results since the true restrictions are imposed. In particular, it produces lowest RMSE for time-varying parameters and factor loadings, and approximates the best the underlying factors. However, note that the maximum likelihood estimation works well in case of low-dimension systems, but it becomes difficult when the number of time-varying parameters is large. In that case, the good initial values for factor loadings are needed.

The bottom line of this simulation experiment is that the simple two-step recursive least square method is a very good starting point to investigate for the presence of commonalities within the time-varying parameters in linear regression and VAR models. A better estimate of the correct number of factors seems to be obtained by applying factor analysis tools on MLE TVPs from the standard TVP model, but there is an initial cost in terms of specification and numerical estimation of the model. Finally, once we have a good idea on the number of common components, all the Factor-TVP coefficients and states should be simultaneously estimated using Kalman filter. I will use exactly the same ordering in the application presented in the next section.

value decomposition and check if the eigenvalues of b_t are inside the unit circle. If not, I scale them such that they are less than one and reproduce the corresponding coefficient matrix.

¹⁰I report the results for rolling window only. The results with expanding window are virtually the same.

5 Empirical evidence on common sources of parameters instability

In this section, I produce new empirical evidence for the existence of low-dimension common sources of parameters instability in various reduced-form models widely used in macroeconomics. A 3-variable TVP-VAR model from Primiceri (2005) is estimated using the three methods discussed above.

5.1 TVP-VAR model

The VAR model contains the annual growth rate of GDP price index, unemployment rate, and 3-month Treasury bill. The data used are quarterly and cover the 1953Q1-2006Q4 period. The lag order is fixed at 2, so the total number of time-varying parameters is 27 (3 constants, 18 VAR coefficients plus 6 covariances). Following notation in Primiceri (2005), the model is

$$y_t = c_t + B_{1,t}y_{t-1} + B_{2,t}y_{t-2} + A_{0,t}^{-1}\Sigma_t\varepsilon_t, \quad (23)$$

where $A_{0,t}^{-1}$ is lower triangular with unit diagonal, and Σ_t is diagonal. Let $x_t = I_n \otimes [1, y'_{t-1}, y'_{t-2}]$, $\beta_t = [c'_t \quad \text{vec}(B'_{1,t}) \quad \text{vec}(B'_{2,t})]'$, and $P_t = A_{0,t}^{-1}\Sigma_t$.¹¹ Then, the model (23) can be written as in (13) with $R_t = P_tP'_t$. Let α_t be a vector of lower triangular random elements in $A_{0,t}^{-1}$, and σ_t contains diagonal elements of Σ_t . The Factor-TVP representation is

$$y_t = x'_t\beta_t + P_t\varepsilon_t \quad (24)$$

$$\begin{bmatrix} \beta_t \\ \alpha_t \\ \log \sigma_t \end{bmatrix} = \lambda f_t + e_t \quad (25)$$

$$f_t = \rho f_{t-1} + v_t \quad (26)$$

A number of studies have documented instabilities within this type of US monetary VAR, see relevant references in Primiceri (2005). At the same time, several researchers have identified changes in some coefficients of Taylor rule approximations of monetary policy. For instance, Boivin (2005) find, in the context of forward-looking Taylor rule with drifting coefficients, that the Federal Reserve's response to expected inflation was strong before 1973 and after 1980, while the response to real activity forecasts have decreased during seventies. Using a quadratic term structure model where the short rate follows a Taylor rule with drifting parameters, Ang et al. (2011) find that the response to inflation have changed substantially over all post-war period, not only under Volcker,

¹¹Let d_t be diagonal matrix containing diagonal elements of P_t . Then, $A_{0,t}^{-1} = P_t d_t^{-1}$ and $\Sigma_t = d_t$.

while the coefficient on output gap presents only small instability over time.

Hence, on one hand we find evidence on instabilities in VAR parameters and stochastic volatilities, and on the other hand literature have identified a small number of structural changes within the monetary policy rule. Suppose that in the theoretical DSGE model the monetary rule changes over time such that parameters on future or past inflation and output are drifting, potentially all coefficients in the VAR representation of observables could be time varying. But the sources of these variations are only few. Thus, commonalities within TVPs of the VAR should be observed. My objective here is precisely to verify this conjecture empirically.

5.1.1 Detecting common parameters' instability

Two-step recursive method As a starting point, I estimate the time-varying parameters and the underlying common sources of instability using the two-step recursive procedure discussed above. Although this method is not the most efficient, the simulation evidence showed that it is able to detect the low-rank variability in TVPs. Due to sampling errors, it tends to suggest more factors than the true number, but it gives a good initial guess and provides starting points for the likelihood-based method. In the first step, I estimate TVPs by recursive OLS where the window is fixed to 10 years.¹²

A necessary condition for factorability is that θ_t presents enough correlation. In this exercise, more than 58% of unique elements in the correlation matrix of θ_t are higher than 0.5, showing that the correlation structure is strong enough to lead to a factor structure. The scree and trace tests are presented in Figure 1. The scree test suggests that there is at least 1 factor. The cumulative product of eigenvalues is also very informative and indicates a small number of factors.¹³ Finally, the trace ratio shows that 2 factors explain almost 80% of the variability in data. Table 2 presents results from the statistical tests for the number of factors. The MAP tests suggest 4 factors, while the Parallel tests estimate 3 and 4 latent components.¹⁴ The Bai-Ng tests suggest 1 to 5 dynamic factors. The next step is to see if the linear approximation of the factor representation of time-varying parameters is reliable. For the sake of space, the results are presented in the online appendix. Overall, the linear hypothesis seems plausible, but adding a polynomial structure would

¹²Different rolling windows' sizes have been fitted, and the results do not vary a lot. Here, I present only results from rolling window approach since it appears to be the most robust.

¹³If the factor structure is true, than all the eigenvalues after the number of factors should be close to zero, and hence, the cumulative product should go to zero. Because of the sampling uncertainty, taking products might be more reliable than looking at eigenvalues only.

¹⁴The first specification of the MAP test works with squared correlation coefficients and the second with power 4. The first two specifications correspond to PCA & Random data generation and PCA & Raw data permutation, while the last two correspond to PAF/Common factor analysis with random data or raw data permutation. The latter and the MAP test tend to overestimate the number of factors in structures with weaker correlation structure. The PCA specifications of Parallel test consider also the variability of the scores. The last two parallel test specifications use adjusted correlation matrices which tend to indicate more factors than warranted (see Buja and Eyuboglu (1992)).

capture some nonlinearities in few parameters. Finally, looking at marginal contribution of common factors on each TVP, I found that the coefficients of the interest rate equation are mostly explained by the first factor. The same factor is important for some parameters in two other VAR equations, together with the second and third common component. The complete results are in the online appendix.

Overall, these results suggest that there are very few common sources of parameter instability. However, given the sampling uncertainty involved in the recursive procedure, it is recommended to estimate the time-varying parameters by an adaptive TVP model and then apply the factor analysis to get a more accurate estimation of the number of factors.

Two-step likelihood method In this section, the TVP-VAR model in (23) is estimated using Bayesian methods as in Primiceri (2005), and all TVPs are collected in θ_t . Then, I apply the same factor analysis as in the previous section. I start discussing results for VAR coefficients only. The stochastic volatilities and covariances are found to be driven by additional orthogonal factors.

The factorability is much stronger than in the case of recursive OLS. Figure 2 shows scree and trace tests. It is obvious that only one factor is sufficient to well approximate all TVPs. Bai and Ng tests also suggest 1 factors, while Parallel test points to 2 components. The linear approximation seems to be plausible, see the corresponding figures in the online appendix. Hence, these tests suggest a linear factor model on $\hat{\theta}_t$ with 1 factor.

TVP-VAR with stochastic volatility I applied the two-step methods to the TVP-VAR with stochastic volatility as in Primiceri (2005). Compared to the previous case where the volatility was supposedly constant, the tests suggest generally the same number of factors. Figure 3 presents the marginal contribution of factors to the total R^2 of TVPs and stochastic volatility parameters. Elements 22 to 24 correspond to coefficients in the contemporaneous relations matrix A_0^{-1} , and the last three elements are the variances of structural shocks. An interesting result is that the VAR coefficients are explained by the first factor, the elements of A_0^{-1} are related to the second factor, while the structural shocks volatilities are related to the third factor.

Hence, the roots of instability in VAR coefficients, that govern the transmission mechanism of shocks, are different from those responsible for movements in covariances (contemporaneous relations in a recursive SVAR) and in variance of shocks. These finding could be of interest for a counterfactual exercise to study the Great Moderation and test the Good Policy (changes in transmission mechanism) versus Good Luck (changes in shocks' variances) hypothesis. The sources of both explanations are small in number and uncorrelated which has significant implications for a counterfactual setup: one can test the explanatory power of each in explaining the fall in variances

of real activity measures. This goes beyond the scope of the actual paper but is a part of ongoing research agenda in Amir-Ahmadi and Stevanovic (2015).

Overall, this exercise shows strong empirical evidence for low-dimension sources of parameters instability in standard monetary VAR model. In particular, it suggests that the VAR coefficients variability is due mostly to one underlying factors. Hence, assuming uncorrelated random walk as dynamic process for VAR drifting coefficients seems not be the most adequate representation in this case. Imposing the factor structure per se is not the only possibility: one could assume reduced rank variation in the vector of TVPs and/or the volatilities. This would reduce sampling uncertainty but does not allow detecting the underlying factors.

5.1.2 Results from the *Factor-TVP VAR*

The next step is to take into account this information explicitly and estimate the Factor-TVP model in one step. The results above show that stochastic volatility is driven by factors other than the sources of time variation within VAR coefficients. The simultaneous maximum likelihood estimation of the Factor-TVP VAR model without stochastic volatility is performed in this section. The Kalman filter is used to calculate the likelihood. The estimated θ_t are plotted in Figure 4. The Table 3 summarizes the results of interest. According to likelihood ratios between 1-, 2- and 3-factor models, not reported here, the 1-factor specification seems to be preferred. Also, the estimation with more than one factor is very unstable and the algorithm tends to find different local maxima. The underlying factor is very persistent, $\hat{\rho}$ close to unity, which is in line with the strong collinearity in θ_t found after the first stage estimation in the two-step likelihood procedure.

The common TVP factor is highly correlated with the unemployment rate, and moderately related to inflation and interest rates. The top panel of Figure 5 plots the VAR series and the estimated factor. To get a more reliable picture, the series are standardized because the factor is much less volatile than the observable series. Finally, the bottom panel of the Figure 5 shows the factor loadings. Note that the errors-in-variable identification is imposed such that the first $(q \times q)$ part of λ is identity matrix. Several coefficients across all three VAR equations load on the common factor.

The stability of the TVP-VAR has always been an issue within this literature [see discussion by Sims and Zha (2006) versus Cogley and Sargent (2005)]. Some suggest to impose the stability condition to hold for each and every time period, while others advocate to let it free and check ex-post if the condition is satisfied in general. Imposing it can be restrictive: (i) from computational point of view, this means discarding potentially a large number of posterior draws in Bayesian setup; (ii) from economic point of view, there might be periods where the economic states are such that the VAR representation of variables is on an explosive path, and discarding them is very

arbitrary. But, these periods of explosiveness must be few. In my setup it is possible to impose the stability restrictions either as nonlinear constraints on loadings and factors or simply by discarding the unstable solutions. However, this is numerically extremely difficult to do within maximum likelihood estimation, hence I do not impose it. Instead, I calculate the maximum eigenvalue of the companion matrix of VAR coefficients for every period and report them in Figure 6. The stability condition is violated if the largest eigenvalue is greater than 1. This happens only during early 1980s that roughly corresponds to the period when the Fed has under Volcker chairmanship. Hence, I conclude that the stability of the implied VAR representation is not an issue in this application.

Evidence using post-crisis data In the previous exercise I used the data up to end of 2006. Recall that the objective of this paper is to establish the empirical evidence on common sources of parameters' instability, following the idea that a small number of structural breaks occur. The recent financial crisis is an example of such a structural change in economic behavior, and considering the most recent data may help in identifying the common shocks. In this section, I repeat the same exercise but with data set updated to 2013Q4 with series from the Federal Reserve of St-Louis FRED data base.

After applying the two-step recursive procedure, the selection procedures suggest more components than with the previous data set: MAP test suggest 2, Parallel 5, while different Bai and Ng (2007) criteria estimate 1, 3, 4, 5 and 6 common shocks. The scree test also suggest an additional factor (not reported here, see figure in the online appendix). Moreover, the marginal contribution of factors to the total R^2 has changed. With the pre-crisis data, the first factor was mostly related to interest rate equation, while using (post-)crisis data produces a picture where the first factor is also important for the unemployment rate equation, and the second factor is mostly related to inflation and unemployment rate equations (graphs are presented in online appendix).

Finally, the simultaneous maximum likelihood estimation of the model containing 2 factors is performed. The summary of estimation results is in the last column of Table 3. Both factors are very persistent with autoregressive coefficients close to unity. The first factor is now highly correlated with the short interest rate, and positively correlated with inflation and unemployment. The second factor is very negatively correlated with unemployment and positively with interest rate, while the correlation coefficient with inflation is only 0.22. Also, the two TVP factors present a correlation coefficient of 0.31.¹⁵ The time-varying VAR coefficients are plotted in Figure 7. For many parameters there is an important change after 2007, indicating the presence of a structural break. The top panel of Figure 8 plots the VAR series and the two factors, while the bottom part presents the factor loadings. Now, contrary to pre-crisis data, a greater number of TVPs load on

¹⁵Recall that the factor model is in errors-in-variable representation, hence the factors dynamics is left unconstrained and they can be correlated.

factors. In particular, all but one inflation equation coefficients seem to be time-varying with the two factors.

5.1.3 Further analysis

TVP factors and hundreds of macroeconomic time series In order to go further in interpretation and significance of the empirical findings in this paper, I face the TVP factors to a large data set containing 141 quarterly macroeconomic and financial series. The full description is available in online appendix. To start, the first panel in Figure 9 shows correlation coefficients, in absolute value, between all series and the TVP factor estimated from the VAR model spanning 1959-2006 period. The most correlated series is, as noted in Table 3, the unemployment rate. The following four most related series are 5 and 10-year Treasury rates, as well as the BAA and AAA corporate bonds, with coefficients close to 0.6. The middle and bottom panels plot correlation coefficients between each of two TVP factors estimated on data ending on 2013Q4. The first factor is now highly correlated, around 0.8, with the yield curve and corporate bonds, as well as with some inflation series. The interpretation of the second factor is different. The highest correlation coefficient, 0.73, is with the utilization capacity in manufacturing, followed by employment indicators (growth rate of employees in: total nonfarm, service-providing, transportation and utilities, and wholesale sectors). Hence, the two sources of variability in the VAR parameters seem to come from the financial markets and from the real sector of economic activity.

TVP factors versus data factors Recently, a number of researchers pointed out that small scale VARs can easily suffer from the lack of information and missing variables problems [see Bernanke, Boivin and Elias (2005) and Forni, Giannone, Lippi, and Reichlin (2009)]. A proposed solution is to augment the VAR with factors estimated from a large number of time series. Let X_t be the macroeconomic panel of 141 series. Suppose it has the following factor structure

$$X_t = \Lambda F_t + \eta_t$$

where F_t consists of K latent factors. These data factors are seen as proxy of macroeconomic fundamentals. Hence, the objective here is to verify the relation between the estimated TVP factors and the data factors. Bai and Ng (2002) IC_{p2} criterion suggests 10 components in F_t . The interpretation of each of them, in terms of R^2 between factors and individual time series, is presented in online appendix. Factors estimated from 1959-2006 data set are such that the first is related to real sector, second to inflation and third to financial markets and unemployment rate. In 1959-2013 data set, \hat{F}_{1t} is clearly a real activity factor while \hat{F}_{2t} is linked to interest rates and inflation series. The interpretation of subsequent factors is less clear.

I start by looking at simple correlations between TVP factors estimated for both data set lengths, (\hat{f}_t^{2006} and \hat{f}_t^{2013}) and data factors (\hat{F}_t^{2006} and \hat{F}_t^{2013}). The TVP factor \hat{f}_t^{2006} is related to the second data factor. Using TVP factors estimated for period 1959-2013 revealed that \hat{f}_{1t}^{2013} is very correlated with \hat{F}_{2t}^{2013} , while \hat{f}_{2t}^{2013} is related to \hat{F}_{1t}^{2013} . To go beyond the simple correlations, I tested if the data factors Granger-cause the TVP factors. In particular, the following multivariate regression is performed:

$$\hat{f}_t^i = \alpha + \rho \hat{f}_{t-1}^i + \beta \hat{F}_{t-1}^i + u_t \quad (27)$$

where $i \in (2006, 2013)$. I test whether $\beta = 0$. In both, 2006 and 2013 samples, the null is largely rejected suggesting that data factors have significant forecasting power for TVP factors. However, quantitatively, the evidence for \hat{f}_{2t}^{2013} is weak since \hat{F}_{t-1} adds a tiny fraction to the adjusted R^2 of regression (27). So overall, this evidence suggest again the TVP factors are linked to output and financial fundamentals of economic activity.

Predictive power of TVP factors Above, I found that TVP factors present some important comovements with output and financial series, as well as with the corresponding common factors. Now, I check the predictive content of the TVP factors for each of 141 macroeconomic time series using Stock and Watson (2002b) diffusion indexes model:

$$z_{t+h} = \alpha_h + \rho_h z_t + \gamma_h \hat{F}_t^i + \beta_h \hat{f}_t^i + u_{t+h} \quad (28)$$

where z_t is one of the macroeconomic series in X_t , F_t^i are the principal components of X_t and f_t^i are the TVP factors, with $i \in (2006, 2013)$. The forecasting horizon, h , varies from 1 to 8 quarters. The idea is to verify if, in sample, the f_t^i have predictive power above the data factors F_t^i .¹⁶ For the sake of space, I present results for the sample ending in 2013Q4 and horizons 1, 4, 6, and 8 quarters. All other results are available in online appendix. The Figure 10 plots the p-values for the null hypothesis $\beta_h = 0$ from the regression (27). It is interesting to see that TVP factors have some predicting power despite the presence of 10 common data factors. Moreover, this forecasting performance rise with the horizon, which can be explained by the fact that data factors \hat{F}_t^i usually improve upon the AR model in short term, but bear less relevant information at longer horizons. In contrast, the TVP factors are very persistent, hence these low frequency movements seem to bring forecasting power at long horizons. For example, when $h = 8$, \hat{f}_t^i seem to be relevant to

¹⁶If the objective was to developp a forecasting model I would do a pseudo out-of-sample analysis as well as a horse rase of several models. But, this would imply re-estimating TVP factors during all out-of-sample periods, which is computationally very difficult. Instead, the idea is simply to verify if the TVP factors have some informational content beyond the past values of the series of interest and data factors. I do it within predictive regression to avoid endogeneity issues. In addition, the regressors of interest in (28) are generated. Some theory has been developed for testing in factor augmented regressions but whit data factors. There is no such results for factors estimated from time-varying parameters. Hence, I present standard p-values as suggestive evidence only.

predict several output series (e.g. industrial production and GDP), investment indicators, many employment series and interest rates. The findings are similar in case of data set 1959-2006. To summarize, the TVP factors are not only correlated with output and financial markets, but they have some explanatory power, even in presence of data factors, for many important macroeconomic series of interest.

5.2 FAVAR model

Since Bernanke, Boivin, and Elias (2005), there is a growing number of literature using the large dimensional factor model to measure the effects of structural shocks in economy. This class of models is of a particular interest here because the number of parameters is very large. Stock and Watson (2007) perform a forecasting exercise using a large dimensional dynamic factor model with time-varying factor loadings, time-varying autoregressive processes for idiosyncratic components and TVP-VAR for latent factors. The results suggest instability in all these parameters. However, given the huge number of TVPs in this empirical application, it is plausible that the underlying variability of these coefficients is of a lower rank. I use the data from Boivin, Giannoni, and Stevanovic (2013), and estimate their factor model by recursive procedure. The model is

$$\begin{aligned} X_t &= \Lambda_t F_t + u_t \\ F_t &= B_t(L) F_{t-1} + e_t \\ \text{Var}(u_t) &= \Psi_t, \quad \text{Var}(e_t) = \Omega_t \end{aligned}$$

where X_t is $N \times 1$, F_t is $K \times 1$, $B(L)$ is of order p , Ψ_t is diagonal. The number of TVPs is $N \times K + N + K^2 \times p + 0.5 \times (K \times (K + 1))$. In their application, $N = 124$, $K = 4$, $p = 4$, which gives 694 TVPs. These are stacked in vector θ_t :

$$\theta_t = [\text{vec}(\Lambda_t)' \quad \text{diag}(\Psi_t)' \quad \text{vec}(B_t(L))' \quad \text{vec}(\text{tril}(\Omega_t))']',$$

where **vec** operator stacks columns of a matrix, **diag** vectorize the diagonal elements of a matrix, and **tril** operator takes only the lower triangular part of a matrix, including its diagonal elements. If one is interested in estimating this model by a likelihood method where the likelihood is obtained with the Kalman filter, even if latent factors were known, there are still 694 states to track. Moreover, it is not very intuitive to rationalize the existence of 694 distinct sources of instability. On the other side, if the factor structure is imposed on θ_t

$$\theta_t = AC_t + \eta_t,$$

with C_t containing $q \ll 694$ elements, there are only q states to filter. Here, I estimate the

TVP-FAVAR model by the recursive two-step procedure with expanding window: the factors are extracted in step 1, then their VAR process is estimated in the second step. Then, I apply the same factor analysis as in the case of TVP-VAR model to the estimated TVPs in $\hat{\theta}_t$.

In the interest of space, I will just summarize the results without showing figures and tables. The scree and trace tests suggest some weaker factor structure than in the VAR case. The MAP and Parallel tests estimate 5 to 10 factors, and the Bai and Ng tests suggest 2 to 4 common shocks. Hence, 4 factors seems to be satisfactory, but one must keep in mind that $\hat{\theta}_t$ contains 694 elements to approximate. As in the VAR case, I find that few factors explain mostly the variations in factor loadings and VAR coefficients, Λ_t and $\text{vec}(B_t(L))$, but less in stochastic volatilities.

When compared to data, the factors have strong predictive power (R^2 between 0.7 and 0.8) for Treasury bonds, especially at long-term maturities, corporate bond yields, some stock market indicators, and the personal saving rate. They are also related, with R^2 between 0.4 and 0.7, to many labor market indicators (employment, unemployment rate, hours, earnings) and some price indicators.

6 Conclusion

The objective of this paper was to detect the presence of the factor structure and test its dimensionality in popular empirical time-varying parameters macroeconomic models. I first showed that structural instability in macroeconomic models is likely to imply time variation in all (or at least a subset of) parameters in reduced-form or estimable models. I discussed how the factor representation of time-varying parameters is obtained, and presented the Factor-TVP that takes into account the factor structure of the parameters. The approach was applied to a standard 3-variable VAR model from Primiceri (2005), and to a factor-augmented VAR (FAVAR) model from Boivin, Giannoni, and Stevanovic (2013).

Overall, I found that parameters' instability in these macroeconomic models is caused by a very few number of common factors. In TVP-VAR application, the underlying factor is very persistent and highly correlated with the unemployment rate, and moderately related to inflation and interest rates. When applied to TVP-VAR model with stochastic volatility, I found that time-variability in VAR coefficients is mostly explained by only one factor, while the stochastic volatility part is explained by two additional factors. In particular, when the structural model is obtained by the Choleski decomposition of the residuals time-varying covariance matrix, the contemporaneous relations matrix parameters were mainly explained by the second factor, while the variances of structural shocks were related to the third factor.

To incorporate the recent financial crisis and to see if there have been common structural breaks,

the same exercise was conducted with the sample updated to 2013Q4. For two-step procedures, the data suggested one more factor, and their interpretation changed. The estimated TVPs presented an important change after 2007, indicating the presence of a structural break. Contrary to pre-crisis data, a majority of TVPs were high correlated with the factors. In particular, all inflation equation coefficients were time-varying with the two factors, and this was also the case for about half of the parameters in other VAR equations.

Comparing TVP factors to a panel of 141 macroeconomic time series and their estimated data factors revealed that (i) the roots of variability in the VAR parameters are likely to come from financial and real sectors of economic activity, (ii) data factors that represent macroeconomic fundamentals Granger-cause the TVP factors, (iii) the TVP factors have significant predictive power, beyond the information contained in data factors, for a great number of output, investment and employment series as well as for term structure of interest rates, (iv) this predictive content increases with the forecasting horizon.

To complete the empirical exercise, the time-varying FAVAR model was estimated by two-step recursive procedure, and 4 dynamic factors were found to drive the dynamics in nearly 700 time-varying parameters. Again, the stochastic volatility coefficients instability had different common factors than the regression coefficients.

Further research is needed to understand the origin and implications of these commonalities. An ongoing research project in Amir-Ahmadi and Stevanovic (2015) studies these issues in a simultaneous Bayesian Factor-TVP framework with stochastic volatility.

References

- [1] Ang, A., Boivin, J., Dong, S. and R. Loo-Kung (2011) “Monetary Policy Shifts and the Term Structure,” *Review of Economic Studies*, vol. 78(2), pages 429-457.
- [2] Amir-Ahmadi, P., and D. Stevanovic (2015), “Common Sources of Instabilities in Macroeconomic Dynamics”, mimeo, ESG-UQAM.
- [3] Bai, J., and S. Ng (2002), “Determining the number of factors in approximate factor models”, *Econometrica* 70:191-221.
- [4] Bai, J., and S. Ng (2007), “Determining the number of primitive shocks”, *Journal of Business and Economic Statistics* 25:1, 52-60.
- [5] Barattieri, A., Eden, M. and D. Stevanovic (2014) “Financial Sector Interconnectedness and Monetary Policy Transmission”, mimeo.
- [6] Bernanke, B.S., J. Boivin and P. Elias (2005), “Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach,” *Quarterly Journal of Economics* 120: 387-422.

- [7] Boivin, J. (2005), “Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data” *Journal of Money, Credit and Banking*, 38(5).
- [8] Boivin, J., and M. Giannoni (2006), “Has Monetary Policy Become More Effective?” *Review of Economic Studies*, 88(3).
- [9] Boivin, J., Giannoni M.P., and D. Stevanović (2013), “Dynamic effects of credit shocks in a data-rich environment,” Staff Reports 615, Federal Reserve Bank of New York.
- [10] Buja, A., and Eyuboglu, N., (1992), “Remarks on parallel analysis,” *Multivariate Behavioral Research*, 27, 509-540
- [11] Canova F. and M. Ciccarelli (2009), “Estimating Multi-country VAR Models”, *International Economic Review*, 50 (3), 929-961.
- [12] Carriero, Clark, and Marcellino (2012), ‘Common Drifting Volatility in Large Bayesian VARs’, working paper
- [13] Cogley T. and T. J. Sargent (2005), “Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.,” *Review of Economic Dynamics*, 8(2), 262-302.
- [14] Edlund, P.O., and H.T. Sjøgaard (1993), “Fixed versus Time-Varying Transfer Functions for Modeling Business Cycles”, *Journal of Forecasting* 12, 345-364.
- [15] Fernandez-Villaverde, J., and J.F. Rubio-Ramirez (2008), “How Structural are Structural Parameters?” *2007 NBER Macroeconomics Annual*, 83-137.
- [16] Fernandez-Villaverde, J., J.F. Rubio-Ramirez, and P. Guerron-Quintana (2010), “Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data,” mimeo, University of Pennsylvania.
- [17] Fernandez-Villaverde, J., J.F. Rubio-Ramirez, T.J. Sargent, and M.W. Watson (2007), “ABCs (and Ds) for Understanding VARS,” *American Economic Review*, 97, 1021-1026.
- [18] Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009), “Opening the Black Box: Identifying Shocks and Propagation Mechanisms in VAR and Factor Models,” *Econometric Theory*, 25, 1319–1347.
- [19] Horn, J.L. (1965), “A Rationale and Test for the Number of Factors in Factor Analysis,” *Psychometrika*, 30, 179-185.
- [20] Inoue, A., and B. Rossi (2011), “Identifying the Sources of Instabilities in Macroeconomic Fluctuations”, *Review of Economics and Statistics*, 93(4), 1186-1204.
- [21] Justiniano A. and G.E. Primiceri (2008), “Time Varying Volatility of Macroeconomic Fluctuations,” *The American Economic Review*, 98(3), pp. 604-641
- [22] Philippon, T. (2012) “Has the US Finance Industry Become Less Efficient?”, *American Economic Review*, forthcoming.
- [23] Pozsar, Z., Adrian, T. Ashcraft, A. and Boesky, H. (2012) “Shadow Banking”, Federal Reserve Bank of New York, Staff Report N. 458.

- [24] Primiceri, G.E. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *The Review of Economic Studies*, 72, 821-852.
- [25] Rao, D.E. (1973). *Linear Statistical Inference and its Applications*, John Wiley & Sons, New York.
- [26] Ravenna, F. (2007), "Vector Autoregressions and Reduced Form Representations of DSGE Models," *Journal of Monetary Economics*, 54:7.
- [27] Ravenna, F. (2010), "The Impact of Inflation Targeting: Testing the Good Luck Hypothesis," mimeo, HEC Montréal.
- [28] Sims, C. A., and T. Zha (2006), "Were There Regime Switches in US Monetary Policy?" *American Economic Review*, 96(1), 54-81.
- [29] Stock, J.H., and M.W. Watson (1996), "Evidence on Structural Instability in Macroeconomic Time Series Relations", *Journal of Business and Economic Statistics* 14:1, 1130.
- [30] Stock, J. H., and M. W. Watson (2002), "Has the Business Cycle Changed and Why?" in *NBER Macroeconomic Annual*, ed. by M. Gertler, and K. Rogoff, MIT Press, Cambridge, MA.
- [31] Stock, J. H., and M. W. Watson (2002b), "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, Vol. 20 No. 2, 147-162.
- [32] Stock, J.H., and M.W. Watson (2003), "Has the Business Cycle Changed? Evidence and Explanations," in *Monetary Policy and Uncertainty*, Federal Reserve Bank of Kansas City, 9-56.
- [33] Stock, J.H. and M.W. Watson (2007), "Forecasting in Dynamic Factor Models Subject to Structural Instability," manuscript, Harvard University.
- [34] Velicer, W. F. (1976), "Determining the number of components from the matrix of partial correlations", *Psychometrika*, 41, 321-327
- [35] Velicer, W. F., Eaton, C. A., and Fava, J. L. (2000), "Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components", in R. D. Goffin and E. Helmes, eds., *Problems and solutions in human assessment*, Boston: Kluwer. pp. 41-71
- [36] Yalcin, I. and Y. Amemiya (2001), "Nonlinear Factor Analysis as a Statistical Method," *Statistical Science*, 16:3, 275-294.

Table 1: Simulation results

Model 1: $T = 300, N = 1, K = 6, q = 1, \sigma_\beta = 0.4$ Factors dynamics: $\rho = 0.95, Q = 0.1$											
Two-step rolling OLS			Two-step MLE				Simultaneous MLE				
Testing number of factors			Testing number of factors								
MAP	1	1	MAP	1	1						
Parallel	1	1	1	Parallel	1	1	1				
Bai-Ng	1	1	1	Bai-Ng	1	2	2				
	1	1	1		1	2	2				
Approximation of factors dynamics			Approximation of factors dynamics				Approximation of factors dynamics				
ρ	0.9986		ρ	0.9852			ρ	0.9568			
Q	0.0038		Q	0.0356			Q	0.0925			
R^2	0.5599		R^2	0.7265			R^2	0.7825			
RMSE: TVPs	0.8625		RMSE: TVPs	0.7494			RMSE: TVPs	0.7088			
RMSE: loadings	0.3177		RMSE: loadings	0.2282			RMSE: loadings	0.1736			
			Log-likelihood	-604.59			Log-likelihood	-583.57			
Model 2: $T = 600, N = 4, K = 4, q = 3, \sigma_\beta = 0.4$ Factors dynamics: $\rho = \text{diag}(0.95, 0.9, 0.85), Q = \text{diag}(0.4)$											
Two-step rolling OLS			Two-step MLE				Simultaneous MLE				
Testing number of factors			Testing number of factors								
MAP	7	7	MAP	3	3						
Parallel	3	3	7	Parallel	3	3	6				
Bai-Ng	1	2	3	Bai-Ng	1	3	3				
	1	2	3		2	3	3				
Approximation of factors dynamics			Approximation of factors dynamics				Approximation of factors dynamics				
ρ	1.0020	0.9981	0.9953	ρ	0.9728	0.9474	0.9309	ρ	0.9487	0.8728	0.8384
Q	0.0027	0.0050	0.0079	Q	0.0615	0.1208	0.1640	Q	0.3745	0.3908	0.4177
average R^2		0.2633		average R^2		0.6832		average R^2		0.8119	
RMSE: TVPs	1.6415			RMSE: TVPs	1.1282			RMSE: TVPs	0.8202		
RMSE: loadings	0.6149			RMSE: loadings	0.3736			RMSE: loadings	0.0724		
				Log-likelihood	-5778.1			Log-likelihood	-5094.1		
Model 3: $T = 300, N = 3, p = 1, q = 1, \sigma_\beta = 0.4$ Factors dynamics: $\rho = \text{diag}(0.95), Q = \text{diag}(0.1)$											
Two-step rolling OLS			Two-step MLE				Simultaneous MLE				
Testing number of factors			Testing number of factors								
MAP	2	2	MAP	1	1						
Parallel	2	2	3	Parallel	3	3	4				
Bai-Ng	1	1	2	Bai-Ng	1	1	1				
	1	2	2		1	1	1				
Approximation of factors dynamics			Approximation of factors dynamics				Approximation of factors dynamics				
ρ	0.9967			ρ	0.9805			ρ	0.9623		
Q	0.0034	0.0201		Q	0.0424			Q	0.0881		
R^2	0.3550			R^2	0.8385			R^2	0.8732		
RMSE: TVPs	0.9039			RMSE: TVPs	0.7809			RMSE: TVPs	0.7172		
RMSE: loadings	0.7531			RMSE: loadings	0.1757			RMSE: loadings	0.0710		
				Log-likelihood	-1411.4			Log-likelihood	-1387.5		

Table 2: Selecting the number of factors in VAR time-varying coefficients

Test	2-step OLS				2-step MLE			
MAP	1	1			8	9		
Parallel	3	3	4	4	2	2	4	4
Bai-Ng	1	3	5		1	1	1	
	2	4	5		1	1	1	

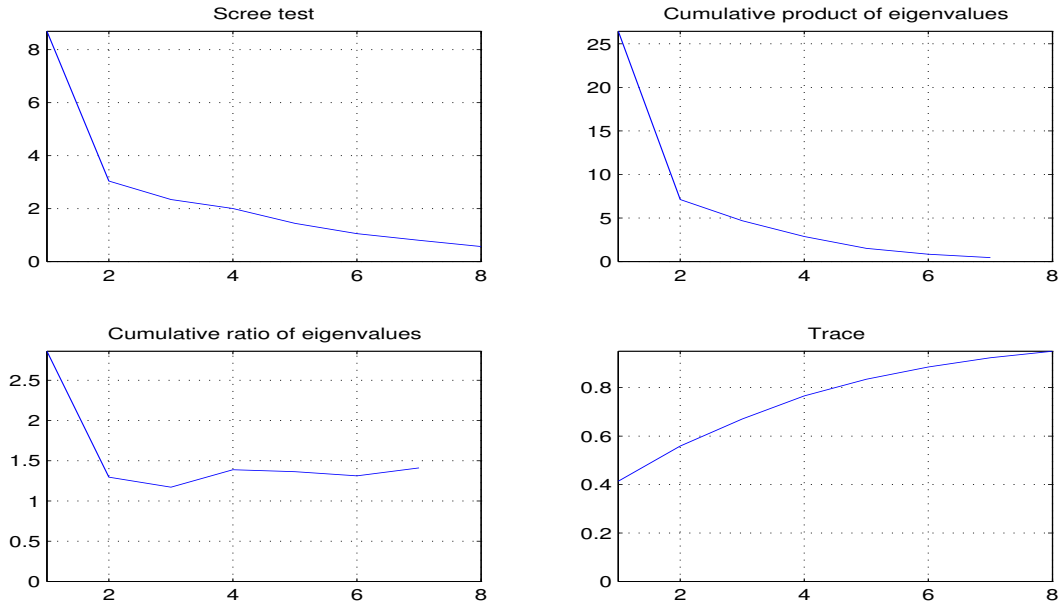
This Table presents the estimation of the number of factors using MAP, Parallel and Bai and Ng (2007) procedures. The procedure are applied to TVPs estimated by recursive least squares (2-step OLS) and to TVPs initially estimated as in Primiceri (2005) (2-step MLE).

Table 3: Summary of ML estimation of Factor-TVP VAR model

	Data up to 2006		Data up to 2013	
Log-likelihood	-233.52		-385.09	
$\hat{\rho}$	0.9996		diag[0.9995 0.9800]	
corr(y, f)	π	0.2430	π	0.4703 0.2248
	ur	0.8419	ur	0.4555 -0.5805
	r	0.3914	r	0.7110 0.5250

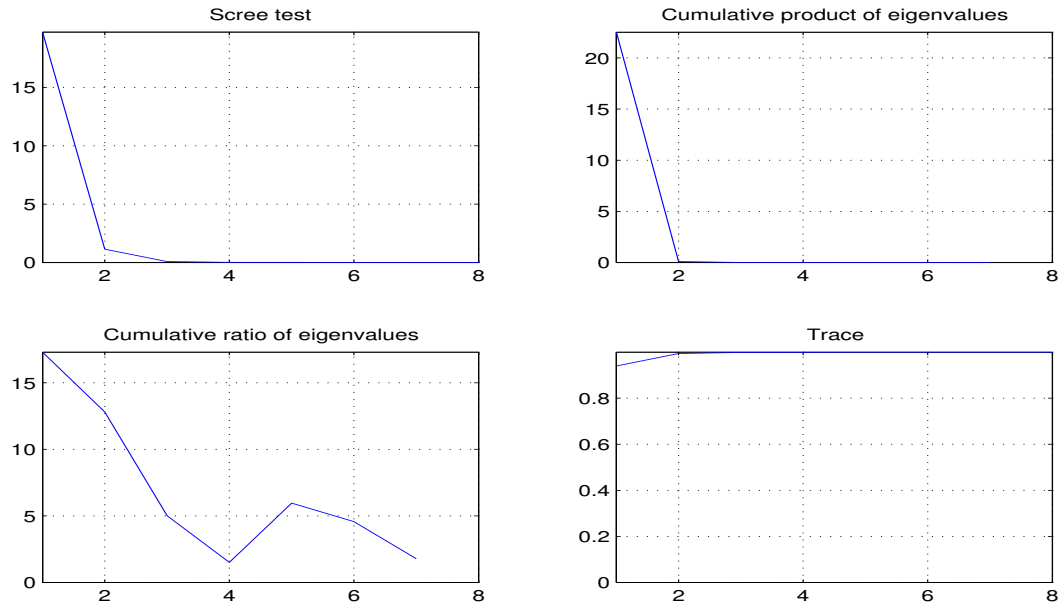
This Table presents some results from the Factor-TVP VAR model estimated in one-step my ML and without stochastic volatility. Column “Data up to 2006” corresponds to estimation on the data set ending in 2006, while “Data up to 2013” refers to data set ending on 2013Q4. In the latter case, there are two common TVp factors. ρ is the autoregressive coefficient in factor’s dynamic process (and diagonal 2×2 matrix in the second column). corr(y, f) contains correlation coefficients between the VAR series and the estimated common TVP factor(s): π stands for inflation rate, ur for unemployment rate and r for the short interest rate.

Figure 1: Scree and trace tests for VAR model TVPs estimated by recursive OLS



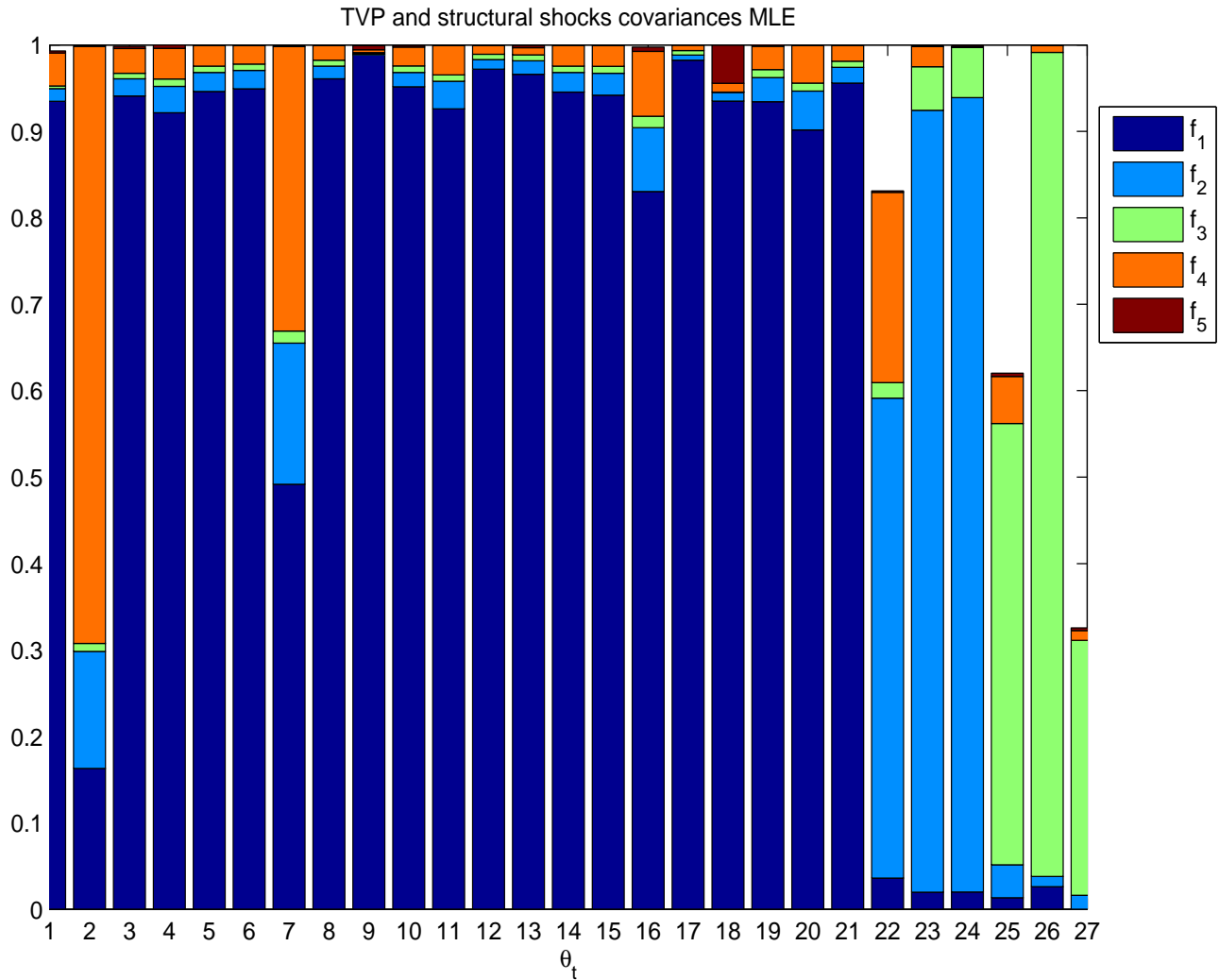
This Figure presents sorted eigenvalues of the correlation matrix of TVPs (scree test), their cumulated product and ratio, as well as the proportion of TVP variances explained by consecutive factors (trace test). TVPs have been estimated by rolling least squares and contain only VAR drifting coefficients.

Figure 2: Scree and trace tests for VAR model TVPs estimated by two-step likelihood method



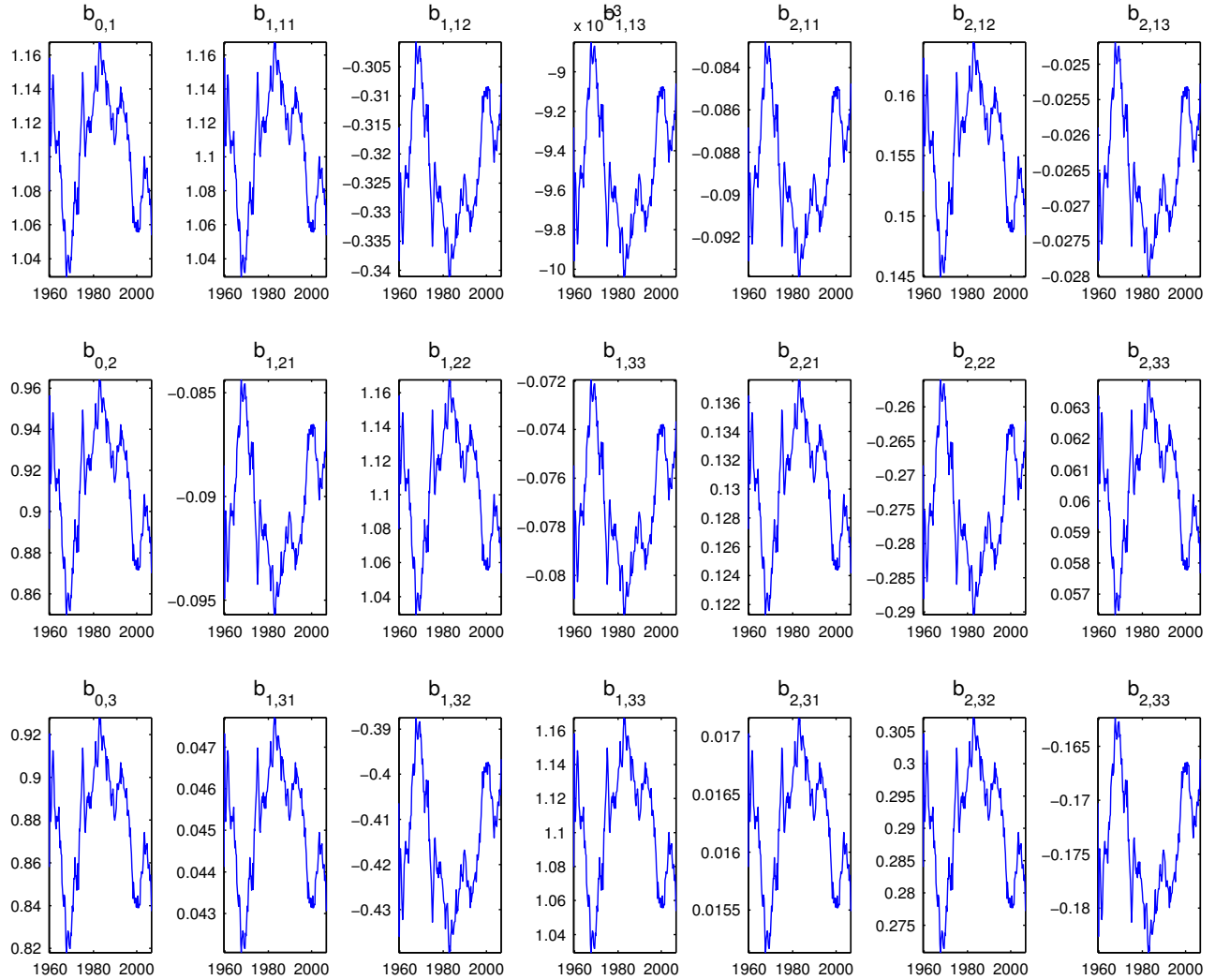
This Figure presents sorted eigenvalues of the correlation matrix of TVPs (scree test), their cumulated product and ratio, as well as the proportion of TVP variances explained by consecutive factors (trace test). TVPs have been first estimated by Bayesian procedure following Primiceri (2005) and contain only VAR drifting coefficients.

Figure 3: Marginal contribution to total R^2 of TVPs and SVs of VAR estimated by two-step likelihood method



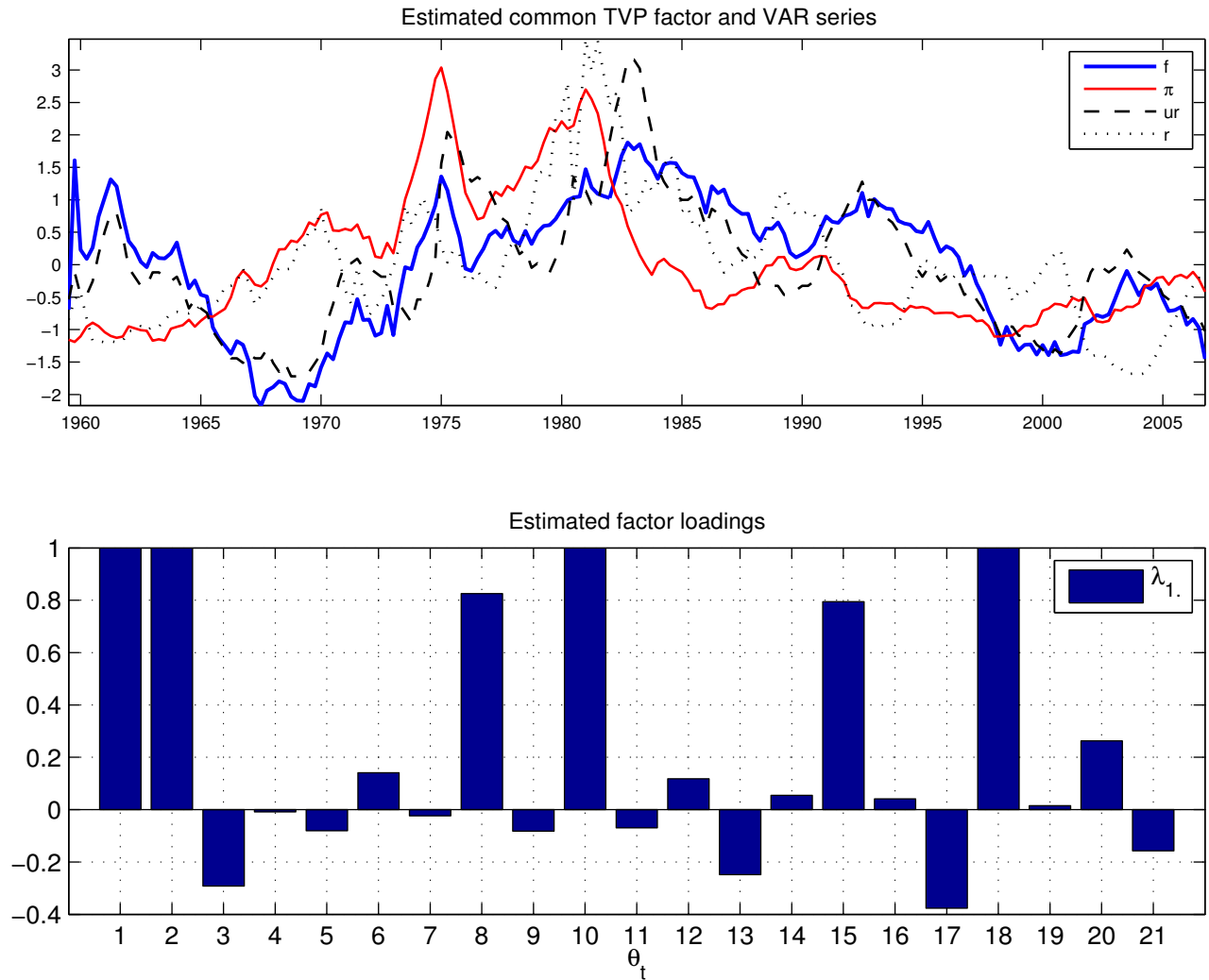
This Figure decompose the contribution of first 5 MLE factors to the variance of each TVP, Covariance state and Stochastic Volatiitiy estimated as in Primiceri (2005). The elements 1-7 represent coefficients from inflation equation, $[b_{0,1}, b_{1,11}, b_{1,12}, \dots b_{2,13}]$, 8-14 those from unemployment rate equation and 15-21 for interest rate. Elements 22-24 are the covariance states and 25-27 are the stochastic volatilities for inflation, unemployment rate and short interest rate equation respectively.

Figure 4: Time-varying parameters from one-step estimation of Factor-TVP VAR



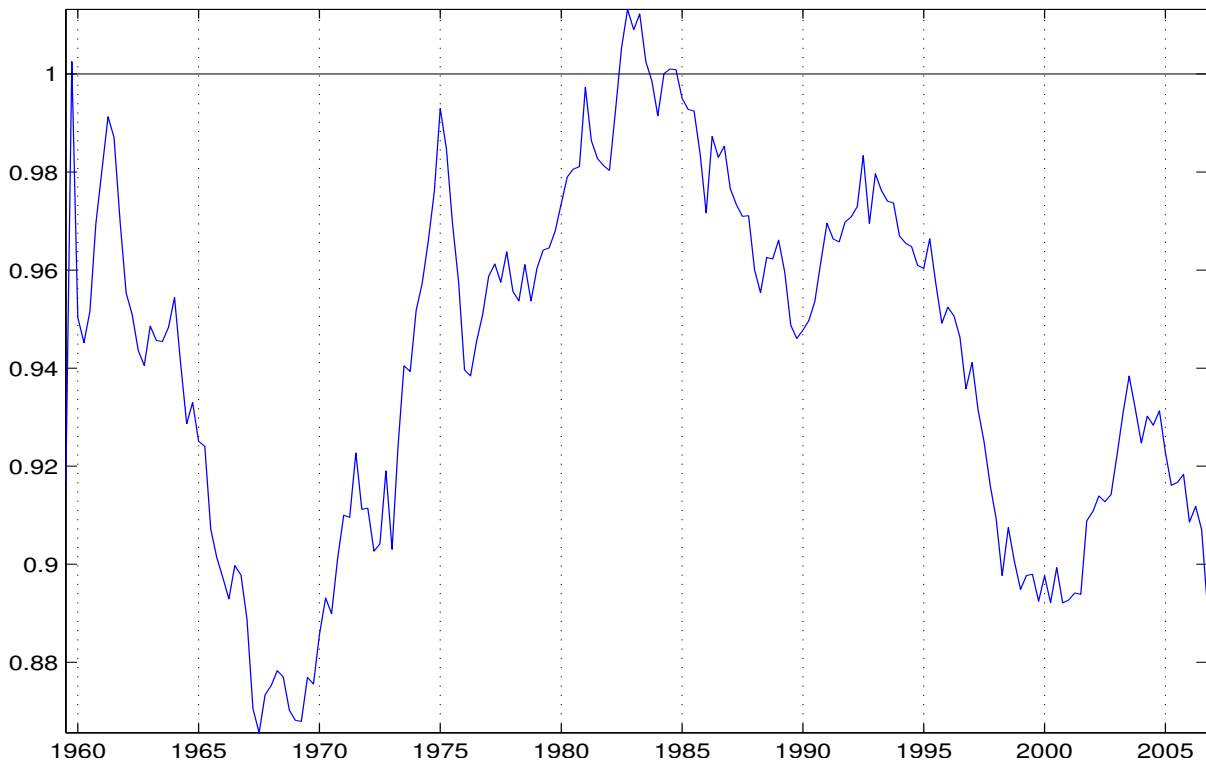
This Figure shows the VAR drifting coefficients calculated as the common component after all the Factor-TVP equations have been estimated simultaneously by maximum likelihood and on the same data as in Primiceri (2005). The top panel shows coefficients from inflation equation, $[b_{0,1}, b_{1,11}, b_{1,12}, \dots, b_{2,13}]$; the middle panel those from unemployment rate equation, $[b_{0,2}, b_{1,21}, b_{1,22}, \dots, b_{2,23}]$; and the bottom for interest rate, $[b_{0,3}, b_{1,31}, b_{1,32}, \dots, b_{2,33}]$.

Figure 5: TVP factor and the loadings from one-step estimation of Factor-TVP VAR



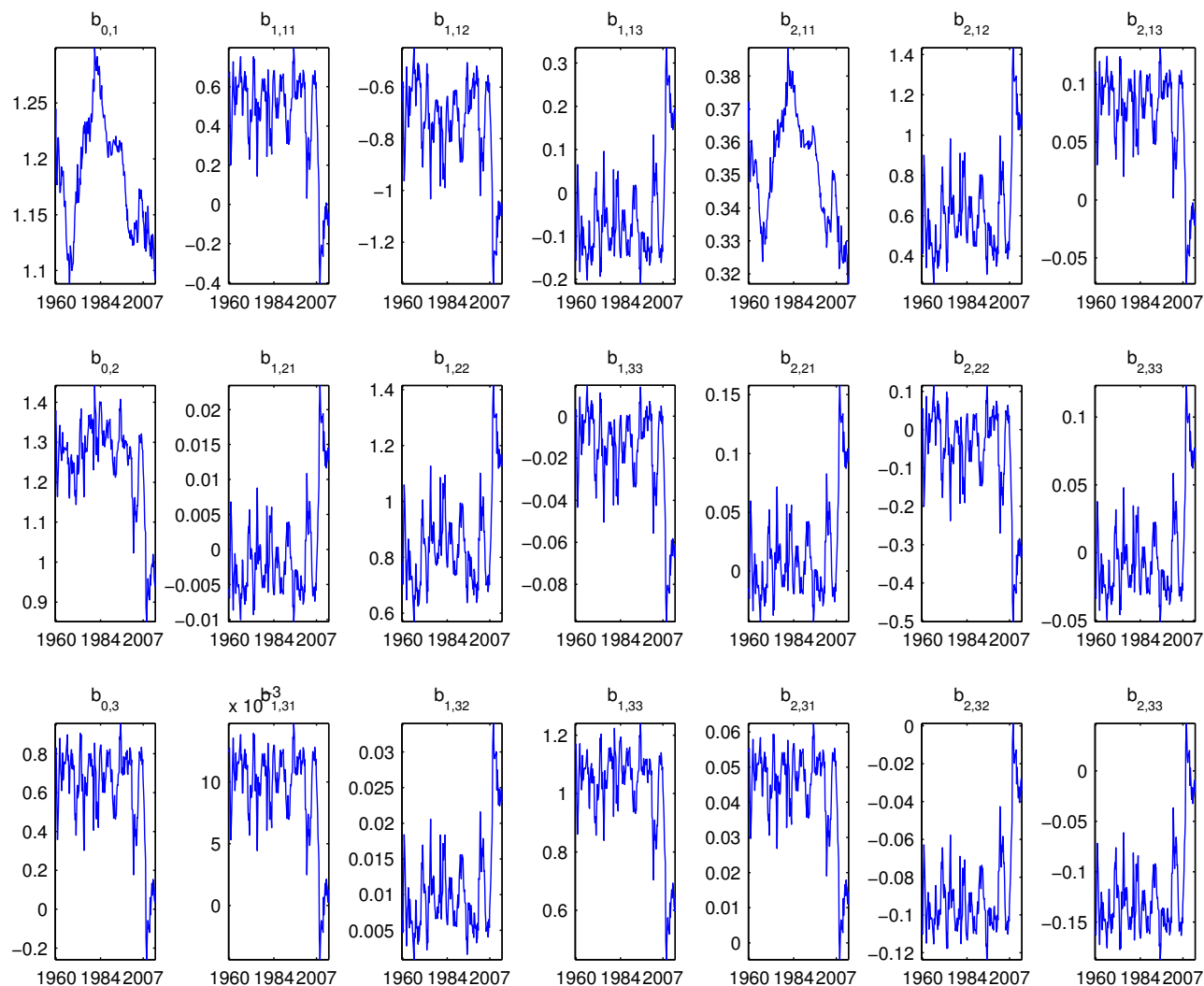
This Figure shows the estimated TVP factor (top panel, together with VAR series) and the corresponding loadings (bottom panel) from the Factor-TVP model estimated simultaneously by maximum likelihood on the same data as in Primiceri (2005). The elements 1-7 represent coefficients from inflation equation, $[b_{0,1}, b_{1,11}, b_{1,12}, \dots, b_{2,13}]$, 8-14 those from unemployment rate equation and 15-21 for interest rate. Elements 22-24 are the covariance states and 25-27 are the stochastic volatilities for inflation, unemployment rate and short interest rate equation respectively.

Figure 6: Stability of Factor-TVP VAR over time



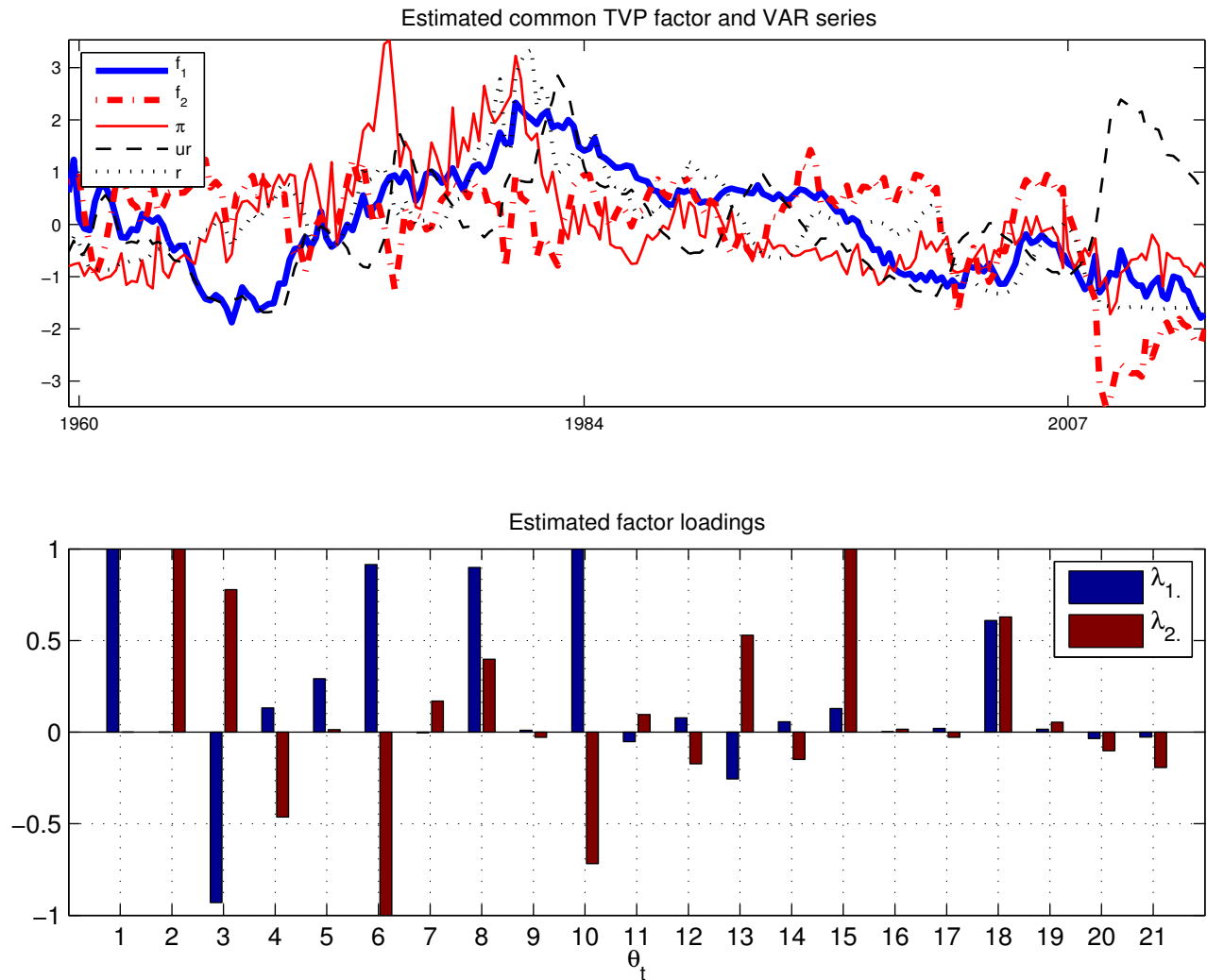
This Figure shows the largest eigenvalue of the companion matrix of the VAR coefficients, at each point of time. The TVPs are calculated as the common component after all the Factor-TVP equations have been estimated simultaneously by maximum likelihood. The stability condition is violated if larger than 1.

Figure 7: Time-varying parameters from one-step estimation of Factor-TVP VAR with post-crisis data



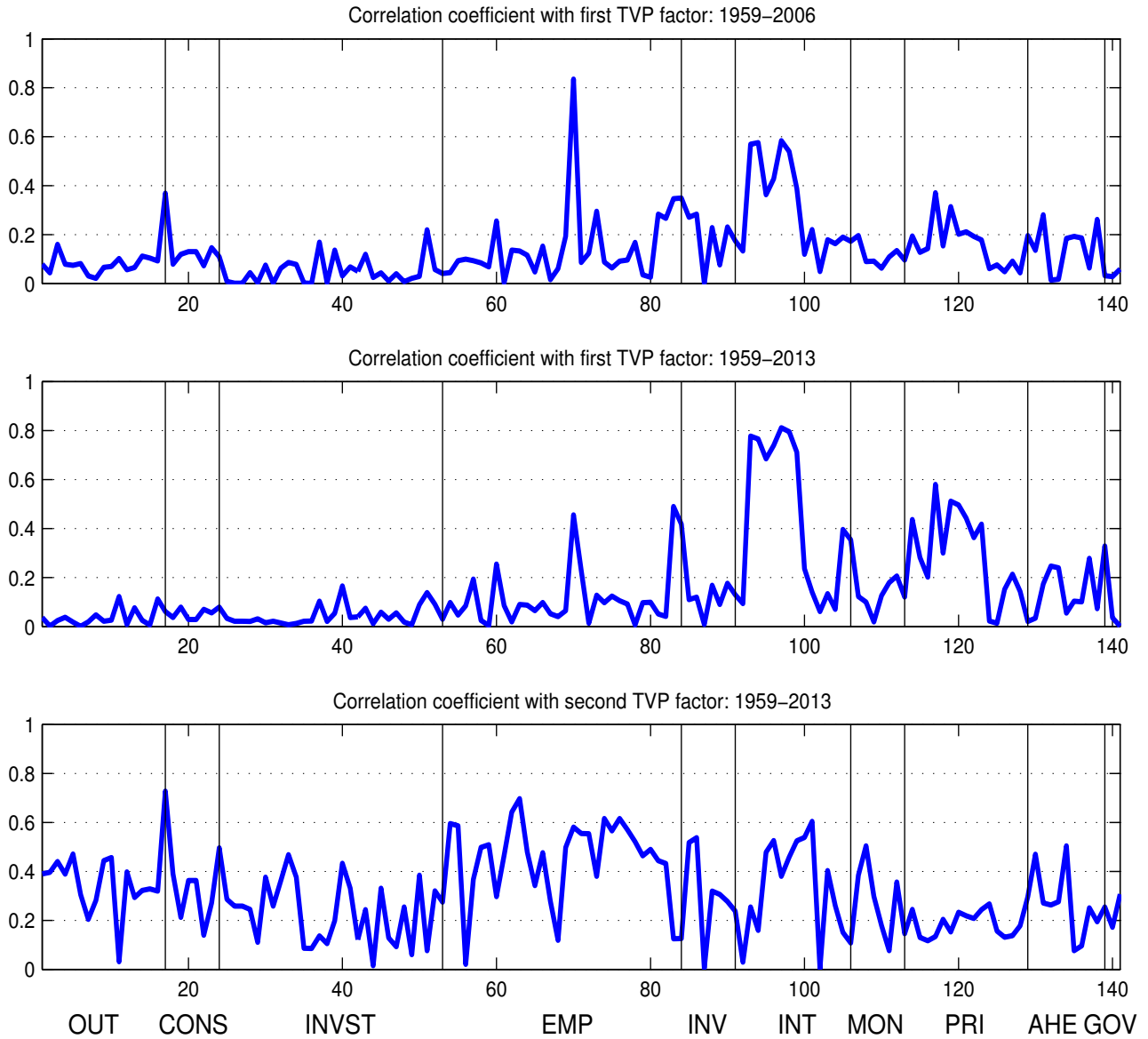
This Figure shows the VAR drifting coefficients calculated as the common component after all the Factor-TVP equations have been estimated simultaneously by maximum likelihood and on the data set updated to 2013Q4. The top panel shows coefficients from inflation equation, $[b_{0,1}, b_{1,11}, b_{1,12}, \dots, b_{2,13}]$; the middle panel those from unemployment rate equation, $[b_{0,2}, b_{1,21}, b_{1,22}, \dots, b_{2,33}]$; and the bottom for interest rate, $[b_{0,3}, b_{1,31}, b_{1,32}, \dots, b_{2,33}]$.

Figure 8: TVP factor and the loadings from one-step estimation of Factor-TVP VAR with post-crisis data



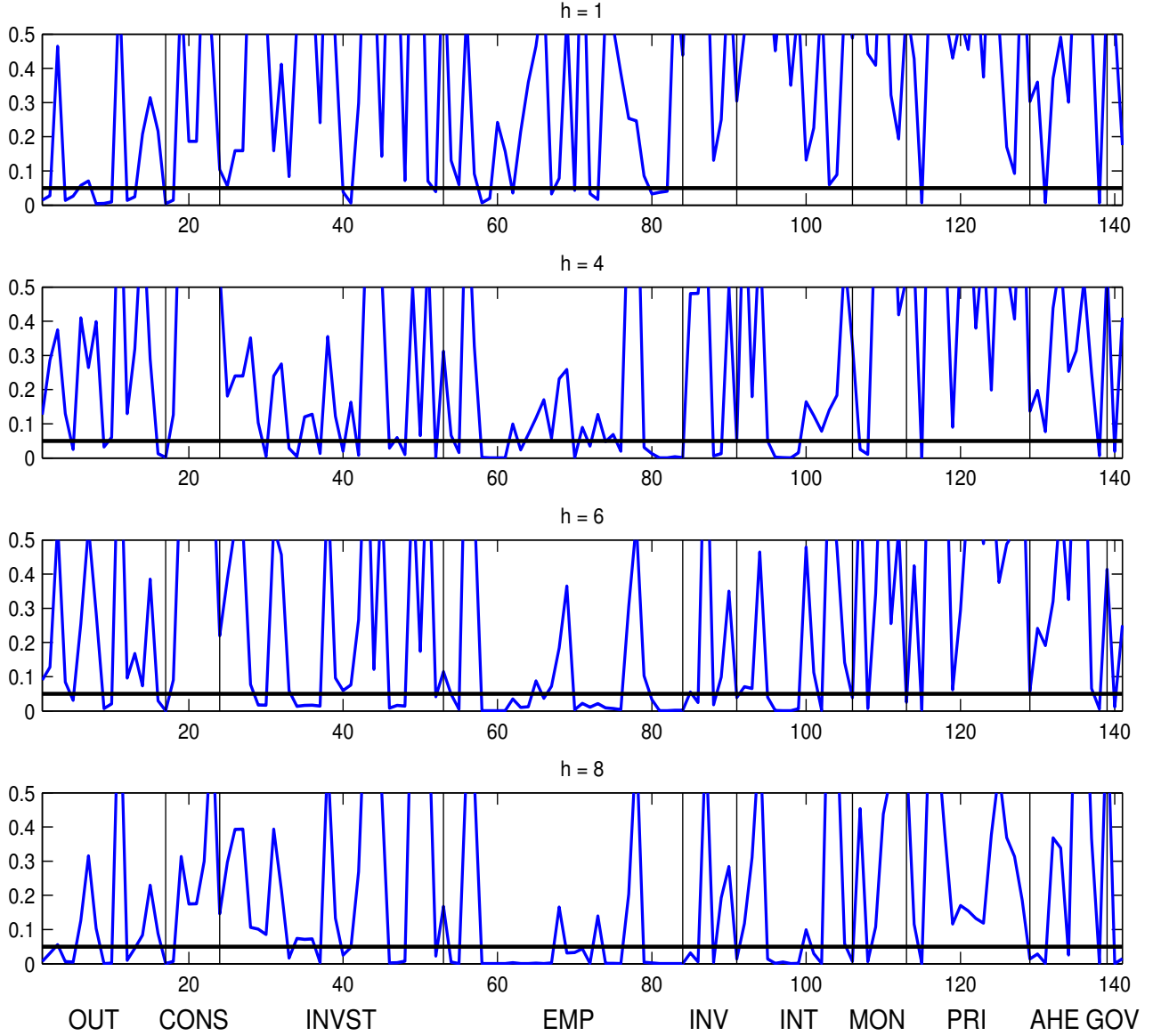
This Figure shows the estimated TVP factor (top panel, together with VAR series) and the corresponding loadings (bottom panel) from the Factor-TVP model estimated simultaneously by maximum likelihood on the updated data set to 2013Q4. The elements 1-7 represent coefficients from inflation equation, $[b_{0,1}, b_{1,11}, b_{1,12}, \dots, b_{2,13}]$, 8-14 those from unemployment rate equation and 15-21 for interest rate. Elements 22-24 are the covariance states and 25-27 are the stochastic volatilities for inflation, unemployment rate and short interest rate equation respectively.

Figure 9: Correlation coefficients between TVP factors and a large number of macro series



This Figure shows contemporaneous correlation coefficients between estimated TVP factors and 141 macroeconomic series grouped according to acronyms below x-axes. The meaning of each acronym is the following. OUT: Real output and income, CONS: Real consumption; INVST: Real investment, EMP: Employment and hours; INV: Inventories; MON: Money and credit quantity aggregates; PRI: Price indexes; AHE: Average hourly earnings and salaries; GOV: Government spending and revenues. The full description of the large data set is available in the online appendix.

Figure 10: Predictive power of TVP factors from data set 1959-2013



This Figure shows the predictive power of two TVP factors for each of 141 series grouped by category. It plots the p -values for null hypothesis $\beta_h = 0$ from the predictive regression (27). The black horizontal line fix the level of 5%. The meaning of each acronym is the following. OUT: Real output and income, CONS: Real consumption; INVST: Real investment, EMP: Employment and hours; INV: Inventories; MON: Money and credit quantity aggregates; PRI: Price indexes; AHE: Average hourly earnings and salaries; GOV: Government spending and revenues.